C. elegans EVI1 proto-oncogene, EGL-43, is necessary for Notch-mediated cell fate specification and regulates cell invasion

Byung Joon Hwang, Alejandro D. Meruelo and Paul W. Sternberg*

During *C. elegans* development, LIN-12 (Notch) signaling specifies the anchor cell (AC) and ventral uterine precursor cell (VU) fates from two equivalent pre-AC/pre-VU cells in the hermaphrodite gonad. Once specified, the AC induces patterned proliferation of vulva via expression of LIN-3 (EGF) and then invades into the vulval epithelium. Although these cellular processes are essential for the proper organogenesis of vulva and appear to be temporally regulated, the mechanisms that coordinate the processes are not well understood. We computationally identified *egl-43* as a gene likely to be expressed in the pre-AC/pre-VU cells and the AC, based on the presence of an enhancer element similar to the one that transcribes *lin-3* in the same cells. Genetic epistasis analyses reveal that *egl-43* acts downstream of or parallel to *lin-12* in AC/VU cell fate specification at an early developmental stage, and functions downstream of fos-1 as well as upstream of *zmp-1* and *him-4* to regulate AC invasion at a later developmental stage.

Characterization of the *egl-43* regulatory region suggests that *egl-43* is a direct target of LIN-12 and HLH-2 (E12/47), which is required for the specification of the VU fate during AC/VU specification. EGL-43 also regulates basement membrane breakdown during AC invasion through a FOS-1-responsive regulatory element that drives EGL-43 expression in the AC and VU cells at the later stage. Thus, *egl-43* integrates temporally distinct upstream regulatory events and helps program cell fate specification and cell invasion.

**KEY WORDS:** EGL-43, HLH-2, FOS-1, LIN-12/Notch, Cell fate specification, Cell invasion

INTRODUCTION

One fundamental issue in development is to understand how cells adopt different fates that lead to their own unique properties. This question becomes more challenging to understand when such cells are adjacent and originate from common ancestral cells. Notch-Delta/Serrate signaling mediates cell-cell interactions of neighboring cells to specify their fates in metazoan development and its dysfunction results in various developmental defects and disease pathologies (Garg et al., 2005; Weng and Aster, 2004; Yoon and Gaiano, 2005).

Specification of the *C. elegans* anchor cell (AC) and ventral uterine precursor (VU) cell fates, initiated by the interaction between ligand LAG-2 (Delta) and receptor LIN-12 (Notch), is a model system to understand the molecular mechanisms through which the Notch signaling pathway specifies the fates of neighboring cells (Lambie and Kimble, 1991; Seydoux and Greenwald, 1989; Wilkinson et al., 1994). In wild-type animals, LIN-12 signaling is necessary for the two somatic gonadal cells [Z1.ppa and Z4.aap; presumptive-VU (pre-VU) cells] to become VU cells. Their sister cells (Z1.aaa and Z4.aaa) have equal potential to become AC or VU cells (Seydoux and Greenwald, 1989; Wilkinson et al., 1994). After AC/VU specification, one of them expresses only LAG-2 and adopts the AC fate, and the other expresses only LIN-12 and becomes a VU cell (Seydoux and Greenwald, 1989; Wilkinson et al., 1994). *lin-12* expression appears to be maintained in the VU cells by a positive-feedback loop that is not well understood. The mechanism that results in the exclusive expression of LAG-2 in the AC appears to involve a transcriptional downregulation of HLH-2, the *C. elegans* ortholog of mammalian E12/E47 and *D. melanogaster* Daughterless (Karp and Greenwald, 2003). The *hlh-2* gene is transcribed in the pre-AC/pre-VU, AC and VU cells, but its encoded protein is absent in the VU cells (Karp and Greenwald, 2003; Karp and Greenwald, 2004). As HLH-2 is required for *lag-2* transcription, the absence of HLH-2 protein appears to be responsible for the elimination of *lag-2* transcription in the VU cells (Karp and Greenwald, 2003).

The distinct cellular properties of the AC and VU cells suggest that they can be used not only to study the mechanisms of Notch-dependent cell fate specification, but also to understand how a cell fate decision is coupled to the later-stage differentiation events. The VU cells undergo four rounds of cell division (Kimble and Hirsh, 1979; Newman et al., 1996), whereas the AC remains a single cell in the gonad and expresses LIN-3, an epidermal growth factor (EGF) that induces proliferation of vulval precursor cells (VPCs) (Hill and Sternberg, 1992). We previously identified a 59 bp *lin-3* enhancer element sufficient to activate *lin-3* transcription solely in the AC. This element consists of two E-boxes and one FTZ-F1 nuclear hormone receptor (NHR) binding site (Hwang and Sternberg, 2004). HLH-2 binds to the E-boxes and is necessary for *lin-3* transcription in the AC (Hwang and Sternberg, 2004). The AC also invades the vulval epithelium by extending a process toward the 1° lineage (P6.p) of vulval cells (Sherwood et al., 2005). The destruction of the underlying basement membranes during AC invasion requires FOS-1, a *C. elegans* homolog of the proto-oncogene Fos, as well as the metalloprotease ZMP-1 and the fibulin HIM-4 (Sherwood et al., 2005).

Genetic screens have identified proteins that modify LIN-12 activity but no physiologically relevant downstream targets involved in AC/VU specification (Greenwald, 2005; Kimble and Simpson, 1997), which might be due to redundancy or pleiotropy of such targets. In this study, we took a bioinformatics-based approach to identify the LIN-12 downstream targets involved in...
AC/VU specification. We first developed a computational program (ClusterSearch) to carry out a genome-wide computational search for genes co-expressed from a common regulatory element consisting of multiple binding sites of sequence-specific DNA-binding proteins. The program identifies orthologous genes sharing a defined regulatory element in the annotated genome sequences of *C. elegans* and *C. briggsae*. Using this program we identified egl-43 as a gene co-expressed with lin-3 during AC/VU specification. EGL-43, which is the *C. elegans* ortholog of the EV11 proto-oncogene that has a sequence-specific DNA-binding activity and acts as a transcriptional repressor, is involved in AC/VU specification as a direct target of LIN-12 and HLL-2. EGL-43 regulates FOS-1 dependent basement membrane breakdown during AC invasion. Characterization of the egl-43 regulatory region indicates that the *C. elegans* orthologs of a set of genes involved in cancer (E12/E47, EGF, EVI1, Fos and Notch) interact to coordinate early-stage AC/VU cell fate decision as well as late-stage events of vulval induction and anchor cell invasion.

### MATERIALS AND METHODS

#### General methods and strains

*C. elegans* strains were maintained at 20°C and cross following standard protocols (Brenner, 1974). Cell anatomy was observed with Nomarski optics, and CFP, GFP and YFP expression were observed using a Hamamatsu ORCA-ER on a Zeiss Axioskop compound microscope. Photographs were taken with a digital camera and Imprint Openlab software, and images were overlaid using Adobe Photoshop CS.

The wild-type strain used in this study is *C. elegans* var. Bristol strain N2. The following variants of N2 were used: unc-119(ed4); syIs107[hsp-40::GFP + pMH86]; syIs57[cdh-3::CFP + pDP#MM016B], unc-119(ed4); syIs107[hsp-40::GFP + pMH86]; syIs57[cdh-3::CFP + pDP#MM016B]; unc-119(ed4); syIs129[hsp-40::GFP + pDP#MM016B]; syIs118[ce-fos-L::YFP + pDP#MM016B]; unc-119(ed4), ppd-1(e1259). The wild-type strain used in this study is *C. elegans* var. Bristol strain N2.

#### Constructs and transgenic lines

The egl-43::YFP translational fusion construct (PR-domain::YFP), which expresses YFP fused with the N-terminal domain of EGL-43 (PR-domain, a known protein-protein interaction domain), was prepared by fusing NLS::YFP in-frame at the second exon of egl-43. 2874 bp of genomic sequence of egl-43 (between 981 bp upstream (5’-AGGAAACTTATAC-3’) and 1893 bp downstream (5’-GCTGATGACGGAGAA-3’) of the translational initiation site) was PCR amplified from N2 genomic DNA and then cloned into the *Stul* and *Ascl* sites in the pPD122.53 (YFP) plasmid. pPD122.53(YFP) was generated by replacing the GFP in pPD122.53 (obtained from Andy Fire, Stanford University, Palo Alto, CA) with GFP and introducing an *Ascl* site 5’ to the nuclear localization sequence. Multiple clones were sequenced to obtain the egl-43::YFP construct containing the wild-type sequence. PCR fusion was used to generate site-directed mutations of E-boxes and LAG-1 binding sites in the 5’ regulatory and first intronic regions of egl-43 (Hobert, 2002). An *MluI* (for mutating E-boxes) or *PstI* (for mutating LAG-1 binding sites) site was introduced during the site-directed mutagenesis but these changes did not affect the egl-43::YFP expression pattern (data not shown). All mutations were confirmed by DNA sequencing. The egl-43 sequence in the egl-43::YFP constructs that contain the wild-type or mutated E-boxes or LAG-1 binding sites is shown in Fig. S1 (see Fig. S1 in the supplementary material).

Transgenic lines for egl-43::YFP constructs were generated using a standard microinjection protocol (Mello et al., 1991). Each YFP construct (25 or 50 µg/ml) was co-injected with pBuescript (170 µg/ml) and myo-2::YFP (5 µg/ml) plasmids. After injection, transgenic animals were identified and maintained by myo-2::YFP expression in the pharynx.

### RESULTS

#### ClusterSearch, a computational program to perform genome-wide searches of evolutionarily conserved regulatory elements

To carry out a genome-wide search for genes co-expressed from a common regulatory element consisting of multiple binding sites of sequence-specific DNA-binding proteins, we developed a computational program (ClusterSearch) that identifies orthologous genes sharing a defined regulatory element in the annotated genome sequences of *C. elegans* and *C. briggsae*. We imported the genomic DNA sequences of *C. elegans* and *C. briggsae* in the FASTA format and their annotation data (gene names, orthology tables, positions of transcriptional initiation sites, exons and introns) in the GFF format from WormBase (http://www.wormbase.org). The program simultaneously scans both genome sequences for the locations of DNA binding

Electrophoretic mobility shift assays with purified HLH-2 and in vitro-translated LAG-1 were performed as described previously (Christensen et al., 1996; Hwang and Sternberg, 2004; Zimmer-Stroh et al., 1994). The sequences of DNA probes and competitors are summarized in Table 1. Chromatin immunoprecipitation (ChiP) with anti-HLH-2 and anti-LAG-3 (SEL-8) antibodies were performed as described previously (Lee et al., 2006).

#### AC invasion analysis

AC invasion was assayed as described previously (Sherwood et al., 2005; Sherwood and Sternberg, 2003). In brief, AC invasion was first evaluated at mid-L3 (granddaughters of P6.p, 4-cell stage) and the L3 molt (great-granddaughters of P6.p, 8-cell stage) using Nomarski optics. Second, the cdh-3::GFP marker, which localizes GFP in the cytoplasm of the AC from late L2 and some vulval cells (from L3 molt), was used to follow the movement and shape of the AC. In wild-type animals, the basolateral side of the AC crosses through a hole in the basement membranes and penetrates between central 1°-fated vulval cells beginning in mid-L3, resulting in the overlap of cytoplasmic GFP signal in the AC with that in the vulval cells at early L4. Third, *spcr::GFP*, a marker for the basement membranes of gonadal and ventral epidermis, was used to directly measure the breakdown of the membranes beneath the AC. In wild-type animals, these membranes are degraded beneath the AC after the L3 molt.

### RNAi analysis


The following variants of N2 were used: unc-119(ed4); syIs107[hsp-40::GFP + pMH86]; syIs57[cdh-3::CFP + pDP#MM016B], unc-119(ed4); syIs129[hsp-40::GFP + pDP#MM016B]; syIs118[ce-fos-L::YFP + pDP#MM016B]; unc-119(ed4), ppd-1(e1259).
defined regulatory elements (binding sites of multiple transcription factors) in a fixed base-pair window. It then summarizes the search results using ontology tables to identify the orthologous genes that contain more than two E-boxes (CANNTG, where N specifies any nucleotide) and one FtzF1 NHR binding site [CA(A|G)(A|T|C)(C|T)] within a 100 base-pair window. Among the 35 orthologs that contain the predicted ACEL-like elements in the 5’-regulatory, first and second intronic regions, we further analyzed the six candidate genes (nhr-22, nhr-25, nhr-91, egl-43, egl-46 and F52F12.6) that have predicted DNA-binding domains. We examined the AC-related phenotypes in worms where RNAi blocked the expression of the six candidate genes; we measured AC/VU specification, AC-specific marker expression and VPC induction. RNAi against two candidates, egl-43 and nhr-25, caused defective phenotypes. nhr-25 RNAi-treated animals showed defective proliferation of VPCs, which could be due to loss of expression in the AC or vulval cells. nhr-25 has been shown to be expressed in the pre-AC/pre-VU, AC and vulval cells (Gissendaner and Studer, 2000). NHR-25 binds to the wild-type FtzF1 NHR binding site, but not to the site containing the lin-3(e1417) mutation, in the ACEL element (Hwang and Sternberg, 2004). egl-43 RNAi caused defects in AC/VU specification and AC invasion, but not in AC-specific lin-3 expression and VPC induction (Fig. 2).

### EGL-43 is involved in LIN-12 (Notch)-dependent AC/VU cell fate specification and in anchor cell invasion

When egl-43 expression was blocked by soaking L1 animals in dsRNA, 78% (63 of 81) of the animals containing AC-specific markers (lin-3::GFP or cdh-3::GFP) (Hwang and Sternberg, 2004;
Pettitt et al., 1996) had at least two ACs expressing GFP at L3 (Fig. 2A-D). Since cdh-3::GFP is expressed in the AC but not in the pre-AC/pre-VU cells, the lack of egl-43 expression appears to change presumptive VU cells into ACs rather than arresting the development of pre-AC/pre-VU cells. There is no predicted secondary target for egl-43 RNAi in the C. elegans genome (http://www.wormbase.org). Furthermore, dsRNAs against different regions of egl-43 (cDNA sequences corresponding to exons 1 to 5, exons 7 to 10, and exons 1 to 10) were equally potent in causing the multiple AC defect (data not shown). In control RNAi experiments with 46 genes (see Materials and methods), only three out of ~900 animals examined showed two ACs, suggesting the high precision of the AC/VU cell fate specification process and the specific role of EGL-43 in this process.

Multiple ACs in C. elegans hermaphrodites have largely been observed in mutant animals (allele n137) in which molecules involved in LIN-12 (Notch) signaling (Greenwald, 2005), or in the early-stage asymmetric division of somatic gonadal cells along the proximal-distal axes (sys-1 to sys-3, gon-14 to gon-16 mutants) (Siegfried et al., 2004), become defective. These sys and gon mutants have high precision of the AC/VU cell fate specification process and the specific role of EGL-43 in this process.

Multiple ACs in C. elegans hermaphrodites have largely been observed in mutant animals (allele n137) in which molecules involved in LIN-12 (Notch) signaling (Greenwald, 2005), or in the early-stage asymmetric division of somatic gonadal cells along the proximal-distal axes (sys-1 to sys-3, gon-14 to gon-16 mutants) (Siegfried et al., 2004), become defective. These sys and gon mutants have a loss or decreased number of distal tip cells (DTCs) owing to the loss of gonad asymmetry (Miskowski et al., 2001; Siegfried et al., 2004; Siegfried and Kimble, 2002). We found that egl-43 RNAi (n=26) did not alter the number of DTCs as compared with the control RNAi (n=23), in which each hermaphrodite contains one DTC in each gonad arm (Fig. 2E,F). Furthermore, the extra ACs are separated in these sys and gon mutants, but are clustered in the mutants involved in the LIN-12 signaling as well as in the animals treated with egl-43 RNAi. Thus, egl-43 appears to work closely with LIN-12 signaling, but not with the Wnt and MAP kinase signaling involved in the asymmetric division of somatic gonadal cells (Siegfried et al., 2004).

Elimination of egl-43 expression reverses the AC-deficient phenotype of lin-12 (gain-of-function) mutant animals; egl-43 RNAi (61%, 44 of 61), but not control RNAi (0 of 26), resulted in lin-12(n137) animals with multiple cdh-3::GFP expressing ACs (Fig. 2G,H). Thus, egl-43 acts downstream of or in parallel to lin-12 during AC/VU specification.

AC invasion was defective in 65% (96 of 147) of animals treated with egl-43 RNAi as judged by the lack of contact between the AC and P6.p descendants visible under Nomarski optics (Fig. 2J,L). All animals (66) treated with control RNAi had attachment of the AC to the underlying P6.p descendants (Fig. 2I,K). When AC invasion was visualized with cytoplasmic GFP (cdh-3::GFP), the invasion defect was evident in the egl-43 RNAi-treated animals: cdh-3::GFP in the AC was distinct from that in vulval cells (vulC, D, E and F), and the AC neither crossed the basement membranes nor penetrated the central 1°-fated vulval cells (Fig. 2I-L). In addition, when the specific breakdown of the basement membranes at the site of AC contact was examined using a basement membrane marker (sparc::GFP), 46% (27 of 58) of animals treated with egl-43 RNAi had intact membranes, but none (0 of 19) treated with control RNAi did (Fig. 2M,N).
Although most animals treated with egl-43 RNAi show defects in both AC/VU specification and AC invasion, the formation of multiple ACs does not appear to be tightly linked to defective AC invasion because egl-43 RNAi sometimes causes a defect in either cell fate specification or the invasion (data not shown).

Furthermore, PR-YFP, a putative dominant-negative form of EGL-43, causes a defect in the invasion but not in AC/VU specification (see Fig. 5).

We observed embryonic-lethality in homozygotes of the null mutant egl-43(tm1802). Both specification and invasion were normal in the egl-43(tm1802)/+ heterozygote and in the homozygotes of the egl-43 mutants (n997 and n1079) (data not shown) in which the deletion of its 3'-regulatory region specifically eliminates EGL-43 expression in the HSN neurons (Garriga et al., 1993).

**EGL-43 is expressed in the AC, VU and DU lineages in the hermaphrodite gonad**

The L2 stage of the hermaphrodite somatic gonad is an amiotic stage in which all 12 somatic gonadal cells stop dividing, the proximal 10 cells rearrange, and one of them, the presumptive AC (pre-AC), moves into the mid-sagittal plane on the ventral surface of the gonad (Hirsh et al., 1976; Kimble and Hirsh, 1979). In our study, we divide the L2 stage into three sub-stages (early, mid and late)...
In the early stage, soon after the pre-AC/pre-VU cells are born from their parents (Z1.pp and Z4.aa) at L1 molt, neither of the Z1.ppp and Z4.aaa cells that have equal potential to become the AC has yet moved to the central position in the gonad. In mid-L2, either the Z1.ppp or the Z4.aaa cell is moving or has moved to the central position on the ventral surface of the gonad. Thus, the AC is visualized at a focal plane different from those of the three VU cells. In late L2 animals, cdh-3::CFP signal can be detected in the AC after it has moved to the central position.

Deletion analysis of the egl-43 genomic region using a YFP reporter identified the 5′ and the first intronic regions as being necessary for its expression in the hermaphrodite somatic gonad.

Fig. 3. Site-directed mutagenesis of E-boxes of the ACEL-like element and LAG-1 binding sites in egl-43. (A) egl-43 gene structure. The ACEL-like element that contains two E-boxes is predicted in the first intron, and eight LAG-1 binding sites are predicted in the 5′ regulatory and the first intronic regions. (B, C) egl-43 (wild type): YFP is expressed (green) at early L2 in the pre-AC/pre-VU cells located in two different focal planes. (D, E) Either the Z1.ppp or Z4.aaa cell moves to the central position on the ventral surface of the gonad to become the AC (green, arrow), which is located (D) in a focal plane different from the two focal planes containing the three VU cells (green). Only one VU focal plane is shown in E. (F, G) The AC begins to express cdh-3::CFP at late L2. The AC (arrow) is appears yellow in F because of the co-localization of cdh-3::CFP (red) and egl-43::YFP (green). (H, I) The egl-43 (mutated E-boxes)::YFP is not expressed in the pre-AC/pre-VU cells, AC and VU cells until late L2. (J, K) egl-43 (mutated E-boxes)::YFP is expressed in the three VU cells (green), and not in the AC (arrow), when the cdh-3::CFP begins to be expressed in the AC (J, red) at late L2. (L-Q) The egl-43 (mutated LAG-1 sites)::YFP is expressed in the pre-AC/pre-VU cells at early L2 (L, M, green). One cell (Z4.aap) expresses YFP very weakly in this animal (L). At mid-L2, YFP expression is retained in the AC (N, in green, arrow), but is lost or decreased in the VU cells (O). At late L2 when the cdh-3::CFP (red) begins to be expressed in the AC (P, arrow), three VU cells regain the YFP signal (Q). The AC expresses both cdh-3::CFP (red) and egl-43 (mutated LAG-1 sites)::YFP (green) at this stage, resulting in the merged yellow colour (P). The egl-43 (mutated LAG-1 sites)::YFP (green) begins to be expressed in two DU cells (yellow arrow) right after the three VU cells regain YFP expression (Q). Only one VU focal plane is shown. For each construct, about 100 animals were examined from six to ten transgenic lines. The pictures represent typical expression patterns at specific stages. The lines include syEx895 and syEx896 for the wild type; syEx891 and syEx892 for the mutated E-boxes; syEx897, syEx898, syEx899 and syEx900 for the mutated LAG-1 binding sites.
which is consistent with a previous report that EGL-43 is expressed in some somatic gonadal cells (Baum et al., 1999). This region contains the predicted ACEL-like element and a cluster of eight LAG-1 binding sites (Fig. 3A). egl-43::YFP expression in the gonad was first detected in the pre-AC/pre-VU cells at early L2 (Fig. 3B,C), and was maintained in their 37 descendants at early L4. egl-43::YFP was also expressed in the AC and VU cells when their cell fates become specified at mid-L2 (Fig. 3D,E). Expression in the two dorsal uterine precursor (DU) cells (Z1.pap and Z4.apa) was first detected at late L2 (data not shown) and was maintained in their descendants. Thus, egl-43 is expressed in the AC, VU and DU lineages in the somatic gonad.

Alternative transcriptional initiations in egl-43 produce two transcripts (α and β) that encode different sizes of in-frame proteins (Garriga et al., 1993). The β form contains only the C-terminal portion of the α form, whose translational initiation was predicted from the ATG codon in the sixth exon. To determine whether these transcripts are differentially transcribed, we generated an egl-43 genomic construct that contains an HA-epitope after the ATG in the first exon, and GFP after the ATG in the sixth exon. Transgenic animals expressing the HA-GFP double tags were then stained with anti-HA antibodies to monitor the expression of the α form, and with anti-GFP antibodies to monitor the expression of both α and β forms. This double labeling experiment indicated that the two transcripts are not differentially expressed, as evidenced by their identical cellular and temporal expression pattern (data not shown).

Multiple regulatory elements are responsible for the temporally dynamic expression of egl-43

We predicted that egl-43 would be co-expressed with lin-3 because they share the same type of enhancer (ACEL), which comprises two E-boxes (‘CACCTG’ form) and one FtzF1 NHR binding site (Hwang and Sternberg, 2004). Site-directed mutagenesis studies indicate that both E-boxes in the ACEL-like element are necessary to express egl-43::YFP in the four pre-AC/pre-VUs and the AC (Figs 3 and 4). These E-boxes are also necessary to express egl-43 in the pre-VU cells undergoing the specification because their mutation eliminates the egl-43::YFP expression until late L2 (Figs 3 and 4).

Since HLH-2 directly binds to the E-boxes in the ACEL-like region in egl-43 (Fig. 6C,D), the requirement of E-boxes for egl-43 expression in the pre-VU cells suggests that egl-43 is involved in the specification of VU fate as a direct target of HLH-2. We could not directly address the involvement of HLH-2 in the egl-43 expression in these pre-AC/pre-VU cells because of the profound effects of the hhl-2 RNAi on the development of the somatic gonad prior to AC/VU specification (Karp and Greenwald, 2003; Karp and Greenwald, 2004).

A cluster of eight predicted LAG-1 binding sites (‘TGGGA’) (Christensen et al., 1996) surrounds the ACEL-like element in egl-43 (Fig. 3A), suggesting a possible interaction between these two regulatory motifs. LAG-1 binds to all of the predicted sites but with different affinities in vitro (Fig. 6A,B). To investigate whether this region of clustered LAG-1 binding sites is a direct target of LIN-12 signaling in vivo, we carried out chromatin immunoprecipitation (ChIP) assays (Fig. 6E). C. elegans Notch signaling is mediated by a ternary complex containing the SEL-8 (LAG-3) transcriptional co-activator, LAG-1, and the intracellular domain of Notch (Petcherski and Kimble, 2000). To detect specific DNA sequences bound to this complex in vivo, we immunoprecipitated SEL-8 in the cellular lysates that contain chromatin DNA fragments from the animals expressing SEL-8::GFP and performed PCR to detect the enrichment of specific DNA sequences in the precipitate (Lee et al., 2006). This ChIP experiment showed that the regions containing LAG-1 binding sites [egl-43 (1) and lip-1], but not the region lacking the LAG-1 binding sites [egl-43 (2)], are enriched in the precipitates with anti-LAG-3 (SEL-8) and anti-GFP antibodies as compared with those with the control IgG and anti-HLH-2 antibodies (Fig. 6E).

Fig. 4. Multiple regulatory elements are required for EGL-43 expression in the AC, DU and VU lineages. (A) Results of the site-directed mutagenesis study and fos-1 RNAi are summarized. White areas represent the absence of egl-43::YFP expression. Green areas represent the presence of YFP expression at specific developmental stages. The areas marked in light green represent decreased egl-43::YFP expression. (B) Model summarizing egl-43::YFP expression patterns in the somatic gonad. The ACEL-like element containing two E-boxes is the major enhancer expressing EGL-43 in the pre-AC/pre-VU cells and in the AC, but not in the VU cells. LAG-1 binding sites are also required for EGL-43 expression in the pre-VU cells when AC/VU cell fates become specified. Unidentified enhancers increase EGL-43 expression in the VU cells and their descendants from late L2, and in the AC from mid-L3.
To determine the effects of the LAG-1 binding sites on egl-43 expression, we mutated the LAG-1 binding sites in the egl-43::YFP construct (see Fig. S1 in the supplementary material) and examined the somatic gonadal expression pattern of individual animals at different stages (Fig. 3L-Q, Fig. 4). Site-directed mutation of the LAG-1 binding sites did not eliminate egl-43::YFP expression in the pre-AC/pre-VU cells and the AC (Fig. 3L-N). However, it eliminated or variably reduced YFP expression in the pre-VU or VU cells at mid-L2 after specification of the AC/VU cell fates (Fig. 3O). Thus, the LAG-1 binding sites appear to be necessary to maintain egl-43 expression in the pre-VU or VU cells during AC/VU specification. We could not directly address the involvement of lag-1 and lin-12 in the egl-43 expression in the VU cells because of the cell fate transformation of VU cells into the ACs upon eliminating lag-1 and lin-12 expression by mutation or RNAi. Since E-boxes in the ACEL-like element are also necessary for the egl-43 expression in the pre-VU cells, interaction between the LAG-1 cluster and the ACEL-like element is likely to allow egl-43 expression in pre-VU and VU cells during the specification process. The interaction between a LAG-1/Su(H) cluster and E-boxes appears to be evolutionarily conserved, as seen in the regulation of Enhancer of split [E(spl)] complex genes during specification of sensory organ precursor cell fates in Drosophila peripheral neurogenesis (Castro et al., 2005).

Although LAG-1 binding sites and E-boxes are necessary for egl-43 expression during AC/VU specification, other unidentified elements appear to support egl-43 expression in the AC from mid-L3 when P6.p daughters begin to divide (Fig. 3C,D), and also in the VU, DU and their descendants from late L2 after AC/VU specification (Figs 3 and 4). This late-stage egl-43 expression in the AC, DU and VU cells was affected neither by mutating E-boxes in the ACEL-like element nor the LAG-1 binding sites, nor by treating animals with lag-1 RNAi (data not shown), which rules out the involvement of cryptic LAG-1 binding sites for the late-stage egl-43 expression in these cells. These late-stage enhancers appear to be responsible for the higher level of egl-43 expression at L3 (Fig. 5O-R), which may be necessary for AC invasion and perhaps also for the patterned proliferation of VU cells. The variable decrease in egl-43 expression caused by mutating the LAG-1 binding sites could be due to the presence of a LAG-1-independent late-stage VU enhancer that strongly drives egl-43 expression from late L2 after AC/VU specification.

The early-stage expression of EGL-43 in somatic gonad appears to be necessary for AC invasion

To determine when EGL-43 is necessary for the AC invasion, we expressed the amino-terminal domain of EGL-43 (PR domain) (Moore et al., 2002) fused with YFP, which allowed us to monitor both expression and phenotypes. We expected this PR domain-YFP protein to be a dominant-negative form of EGL-43 because the PR domain regulates the oligomerization of the EVI1 protein (Nitta et al., 2005), and EVI1 protein variants lacking the PR domain can cause leukemia (Fears et al., 1996; Morishita et al., 1988). AC invasion was defective (60%, 26 of 43) when PR-YFP expression was driven by the wild-type regulatory region of egl-43 (Fig. 5E). This construct expresses the fusion protein in the VU and DU lineages as well as in the AC from early L2 (Fig. 3, Fig. 4, Fig. 5A,B). By contrast, AC invasion was normal (50 out of 50) when the PR-YFP protein was expressed using the region containing mutated E-boxes (Fig. 5F) that eliminates the early expression in the AC and VU lineage (Fig. 4, Fig. 5C,D). Thus, the early-stage expression of EGL-43 in these somatic gonadal cells appears to be necessary for...
invasion. The invasion defect caused by expressing PR-YFP (Fig. 5) appears identical to that caused by egl-43 RNAi (Fig. 2). Since HLH-2 is required for inducing VPCs, we could not directly examine the role of HLH-2 in AC invasion because this process requires the presence of the vulva (Sherwood and Sternberg, 2003). However, both E-boxes in egl-43 are necessary for its expression in the AC and for AC invasion.

**egl-43 is a late-stage-specific downstream target of FOS-1 during AC invasion**

To identify genes that interact with egl-43 during AC invasion, we analyzed the expression of genes previously implicated in the invasion (Sherwood et al., 2005); more than 30 animals were examined for each transgene. egl-43 RNAi did not greatly alter AC expression of cdh-3::GFP and fos-1::YFP (Fig. 5K,L), but severely reduced AC expression of zmp-1 and him-4 (hemicentin) (Fig. 5G-J). Since egl-43 appears to be upstream of zmp-1 and him-4, but not of fos-1, we tested whether fos-1 is an upstream of egl-43. Indeed, fos-1 RNAi eliminated egl-43::YFP expression in the DU and VU descendants as well as in the AC (Fig. 5M,N), suggesting that fos-1 is an upstream regulator of egl-43 in these gonadal cells. Since unknown late-stage enhancers are responsible for upregulating egl-43 expression in these cells (Figs 3 and 4), we tested whether egl-43 expression is sensitive to fos-1 RNAi at specific stages and found that fos-1 RNAi does not decrease egl-43 expression at L2 during AC/VU specification (Fig. 4A, Fig. 5O,P). This finding is consistent with the fact that fos-1 is not involved in AC/VU specification (Sherwood et al., 2005), and also suggests that egl-43 is a late-stage-specific (from L3) downstream target of FOS-1 during AC invasion.

**DISCUSSION**

In this paper, we computationally identified egl-43 as a gene co-expressed with lin-3 in the pre-AC/pre-VU cells and in the AC. In addition to this predicted regulatory element in egl-43, additional elements including a cluster of LAG-1 binding sites and a FOS-1 responsive element drive egl-43 expression for proper AC/VU specification and AC invasion. Although there are many possible mechanisms through which EGL-43 regulates two different cellular events, we favor the model that different expression levels of EGL-43 could regulate the expression of different sets of genes. For
AC/VU specification, a low level of EGL-43 expression is supported by HLH-2, LAG-1 and LIN-12, and for AC invasion, a higher level of EGL-43 expression is provided by the addition of FOS-1 and other unidentified factors. However, an equally plausible model is that different sets of genes interact with egl-43 during two different cellular processes.

We observed egl-43::YFP expression in the Z1.pp and Z4.aa cells in some transgenic lines, suggesting its weak expression in these cells. By contrast, YFP is consistently expressed in their daughter cells in all transgenic lines of the egl-43 (wild-type)::YFP construct examined. It is possible that we could not detect a difference in egl-43::YFP expression between the pre-AC and the pre-VU cells because of the long half-life (longer than 12 hours) of YFP protein and the presence of multiple temporally active regulatory elements that support the dynamic egl-43 expression in these cells and their descendants at L2 and L3 (Figs 3 and 4). Since E-boxes are required for egl-43 expression in these cells during AC/VU specification (Figs 3 and 4), and because the pre-AC contains higher amounts of HLH-2 protein than do pre-VU cells (Karp and Greenwald, 2003; Karp and Greenwald, 2004), the pre-AC could contain transiently higher amounts of EGL-43 than pre-VU cells. However, it is also possible that the cluster of LAG-1 binding sites is responsible for higher EGL-43 expression in the pre-VU cells than in the pre-AC. Unfortunately, the long half-life of YFP prevented us from measuring any difference in egl-43 transcription during the specification process. Destabilized YFP expression derived by the egl-43 regulatory region was not strong enough for us to detect (data not shown). However, considering the HLH-2 expression pattern (its specific disappearance in the VU cells), and the nature of the LAG-1 cluster that expresses genes in the pre-VU or VU cells during the specification process, we speculate that EGL-43 is differentially expressed between pre-AC and pre-VU cells during AC/VU specification.

Other than basic helix-loop-helix (bHLH) proteins of the Hairy/E(spl) family (Bailey and Posakony, 1995; Jarriault et al., 1995; Neves and Priess, 2005), the zinc-finger protein EGL-43 represents the only known direct downstream transcription factor of Notch involved in cell fate specification. Elimination of egl-43, hlh-2, lag-1 or lin-12 expression results in extra ACs, suggesting that these genes are necessary to specify the VU cell fate, and failure of this specification results in the pre-VU cells adopting the default AC fate (Seydoux and Greenwald, 1989). This model is supported by our finding that egl-43 is a downstream target of HLH-2 and LAG-1/LIN-12 in the pre-VU cells, as indicated by genetic epistasis analysis, characterization of the regulatory elements in egl-43, and DNA-binding experiments.

Although EGL-43 is necessary for both the migration of HSN neurons and AC invasion, most of the other genes involved in HSN migration do not appear to be involved in AC invasion (Garriga and Stern, 1994). RNAi against 14 out of 16 genes involved in HSN migration did not cause a defect in AC invasion (data not shown), suggesting that the AC invasion mechanism is different from the mechanism through which the neuronal cell body migrates at long distance.

Two lines of evidence suggest that EGL-43 is an indirect downstream target of FOS-1. First, DNA-binding consensus sequences of mammalian Fos proto-oncogenes (‘TGACTCA’) do not exist in the regulatory region of egl-43, which is sufficient to express egl-43::YFP in the somatic gonadal cells in a FOS-1-dependent manner. Second, ChIP analysis to measure the binding of FOS-1::GFP protein on the egl-43 regulatory region could not identify any direct physical interaction between FOS-1::GFP and the egl-43 regulatory region in the evl-5 animals in which the defects of fertility and AC invasion were rescued by the expression of fos-1::GFP (data not shown). Since EVI1, the mammalian ortholog of EGL-43, is a transcriptional repressor (Bartholomew et al., 1997; Perkins et al., 1991), we speculate that egl-43 is an indirect upstream factor of zmp-1 and him-4. Thus, along with the identification of the downstream factor(s) of FOS-1, which activates egl-43 expression during AC invasion, identification of the EGL-43 direct downstream factor(s) that activates zmp-1 and him-4 in the AC will be an important step to understanding the regulatory network that regulates AC invasion.

HLH-2 is required for LAG-2 expression during AC/VU specification (Karp and Greenwald, 2003), and LIN-3 expression is required for vulval induction (Hwang and Stern, 2004). We show that HLH-2 is necessary for EGL-43 expression, along with LIN-12/LAG-1 during AC/VU specification and with FOS-1 during AC invasion (Fig. 7). Combining these results, we conclude that two cellular events following AC/VU specification, induction of vulval precursor cell proliferation and anchor cell invasion, are initiated by factors involved in AC/VU specification, HLH-2 and EGL-43.

We thank C. Van Buskirk, G. B. Choi, J. Fernandes, E. Hallem, M. Kato, C. Robinson, and G. Schindelman for carefully reading the manuscript; J. DeModena for Ab staining; S. Gharib for help with the RNAi screen; S. Mitani at Tokyo Women’s Medical University for the egl-43(tm1802) strain; J. Kimble at the University of Wisconsin-Madison and Michael Krause at NIH for LAG-3 (SEL-8) and HLH-2 antibodies; M. H. Lee at the University of Wisconsin-Madison for helping with the ChIP assay; and B. Gupta, H.-M. Muller, A. Saldanha, E. Schwarz, and L. Warren for discussions about ClusterSearch. P.W.S. is an investigator of the Howard Hughes Medical Institute. B.J.H. is supported by a NHGRI/NIH Genome Scholar Development and Faculty Transition Award (K22HG02907-02). The development of ClusterSearch was supported in part by the Caltech/SURF award from Dr Ashley to A.D.M.
Supplementary material
Supplementary material for this article is available at
http://dev.biologists.org/cgi/content/full/134/4/669

References


**Stu I/Nru I site**

```
aggaacttattacagacatcgcgcgacaaaatataggggcatgttttctctactagaaatgctcaatagattctaaattt
tagtaattctatgatccctttactggtaaccccatagttctcttctttaagaattttctatattaaaatagtctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
Stu I/Nru I site –

aggaaacttattacagacattcagcgcacaaataaaacggctgcttctactagaattttattttagtattttttttcaatagatattgaagatagttttgggtttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt